Journal of Geophysical Research: Space Physics

RESEARCH ARTICLE
10.1029/2018JA025892

Special Section:
Long-term changes and Trends in the Middle and Upper Atmosphere

On Long-Term SABER CO_2 Trends and Effects Due to Nonuniform Space and Time Sampling

Ladislav Rezac[1], Jia Yue[2], Jian Yongxiao[2], James M. Russell III[2], Rolando Garcia[3], Manuel López-Puertas[4], and Martin G. Mlynczak[5]

[1]Max-Planck-Institut für Sonnensystemforschung, Germany, [2]Center for Atmospheric Science, Hampton University, Hampton, VA, USA, [3]National Center for Atmospheric Research, Boulder, CO, USA, [4]Instituto de Astrofisica de Andalucía, CSIC, Granada, Spain, [5]NASA Langley Research Center, Hampton, VA, USA

Key Points:
- The Level2C SABER daytime CO_2 data set has a nonuniform spatial and temporal sampling, and also shows strong local time variation
- A forward modeling study using SD-WACCM outputs shows that time averaging data into 60-day bins partially compensates for this effect
- The reanalysis of the SABER CO_2 yields a nearly constant up to 90 km, however reaching 8% per decade at 105 km

Correspondence to:
L. Rezac,
rezac@mps.mpg.de

Citation:
Rezac, L., Yue, J., Yongxiao, J., Russell, J. M. III, Garcia, R., López-Puertas, M., & Mlynczak, M. G. (2018). On long-term SABER CO_2 trends and effects due to nonuniform space and time sampling. *Journal of Geophysical Research: Space Physics*, *123*, 7958–7967. https://doi.org/10.1029/2018JA025892

Received 13 JUL 2018
Accepted 2 SEP 2018
Accepted article online 6 SEP 2018
Published online 27 SEP 2018

Abstract The Sounding of the Atmosphere using Broadband Emission Radiometry (SABER) instrument on board the TIMED satellite has been continuously operating for more than 16 years, since 2002, monitoring the CO_2 concentration on nearly a global scale in the middle and upper atmosphere (from 65 km up to 110 km). A recent reanalysis (Qian et al., 2017, https://doi.org/10.1002/2016JA023825) concluded that different deseasonalizing methodologies may have a strong impact on long-term trend analysis, ultimately yielding different altitude profiles of the global mean CO_2 trend. In this work, we aim to understand how the nonuniform spatial and temporal sampling inherent in the SABER CO_2 data set affects the determination of the long-term trends. In addition, our goal is to disentangle reported differences in SABER CO_2 trends due to different time averaging windows and methodologies used for trend estimation. The Whole Atmosphere Community Climate Model is used for synthetic studies of the time series. We demonstrate that, due to the time varying data gaps and nonuniform sampling of local times, different time binning of the SABER CO_2 data may indeed bias the long-term trend estimation. We show and discuss how the 60-day averaging reduces the bias in relative trends. We also conclude that different deseasonalizing methodologies (averaged over the same temporal bins) yield negligible differences on the trend determination. Taking this into account the global mean CO_2 relative trend does not deviate statistically from the tropospheric value below 1×10^{-3} mb (90 km). Above about 90 km, there is a positive slope in the global CO_2 trend profile, but with substantially reduced magnitude for 60-day binned data.

1. Introduction

There is strong evidence for a negative trend in global mean thermospheric (≈400 km) density provided by monitoring satellite drag covering nearly five solar cycles (Emmert, 2009, 2015). Similarly, the day-night averaged ionospheric ion temperature data measured at midlatitude observations at Millstone Hill indicate a decreasing trend in the last four decades (Zhang & Holt, 2013). Among several contributing factors, the radiative cooling effect of CO_2 in the upper atmosphere is generally accepted to be one of the most important drivers of the global long-term changes in the mesosphere-thermosphere-ionosphere system (Laštovička et al., 2014; Qian et al., 2013; Roble & Dickinson, 1989).

Tropospheric data (near surface) observations spanning now more than five decades reveal a clear positive trend of about 5.8% per decade (www.esrl.noaa.gov/gmd/ccgg/trends/). Because CO_2 is chemically inert in the Earth's atmosphere and due to the strong eddy diffusion, the mean CO_2 volume mixing ratio, VMR, altitude profile remains constant well into the middle mesosphere. There, it begins to decrease with increasing height due to molecular diffusive separation and photolysis (Garcia et al., 2014). The height where the CO_2 VMR departs from a well-mixed value (typically just above 80 km) varies in latitude and season as recently revealed by three independent satellite data sets (López-Puertas et al., 2017; Rezac, Jian, et al., 2015). The data combined with models indicate that both, the large-scale atmospheric circulation as well as external solar forcing, influence the upper atmospheric CO_2 spatiotemporal distribution. Therefore, the important question arises regarding which processes characterize and determine the secular long-term changes in the upper atmospheric CO_2? Is the secular long-term trend in the mesosphere and lower thermosphere (MLT) identical to that measured in the troposphere?

The first evaluation of the long-term trend of CO_2 in the MLT, performed by Emmert et al. (2012), revealed a secular trend of 8–10% per decade at 100 km. The mean CO_2 trend exhibited a tendency to increase with altitude between 90 and 100 km. These results relied on the solar occultation data between 2004 and 2012 from the Atmospheric Chemistry Experiment Fourier Transform Spectrometer (ACE-FTS) on board the SCISCAT-1 satellite (Beagley et al., 2010; Bernath et al., 2005). Garcia et al. (2016) independently analyzed the ACE data and reached a similar conclusion to that of Emmert et al. (2012) and Yue et al. (2015). Daytime measurements made by the Sounding of the Atmosphere using Broadband Emission Radiometry (SABER) simultaneously inverted for temperature and CO_2 VMR (Rezac, Kutepov, et al., 2015) and analyzed by Yue et al. (2015) for a period of 13 years (2002–2014) within latitudes ±54°. That study also yielded results consistent with the large linear trend in the CO_2 VMR. The trend increased from about 5% per decade at 80 km to 12% per decade at 110 km. Therefore, within their uncertainties, the ACE-FTS and SABER derived trends were consistent with each other. The SABER data also allowed the calculation of the latitude distribution of the CO_2 trend revealing a hemispheric asymmetry with the northern hemispheric trend (>10%) being nearly twice that of the southern hemisphere in the 100 km altitude region. However, a global circulation model developed at the National Center for Atmospheric Research (NCAR), the Specified Dynamics Whole Atmosphere Community Climate Model (SD-WACCM) did not show such a trend distribution or any increase in the trend above 90 km altitude.

In a most recent study, Qian et al. (2017) have reanalyzed the same SABER and ACE-FTS CO_2 data for trends applying different *deseasonalizing* procedures. Their results indicate that the MLT trends are nearly constant with altitude and largely consistent with the tropospheric value of 6% per decade. Qian et al. (2017) argued that the method of *mean of the residuals*, which yields a smaller trend, is more appropriate because it is essentially independent of the running average bin size (deseasonalizing bin sizes of 7.6, 30.4, and 60.8 days were tested). In the mean of the residuals method the time series are formed by averaging the residuals (60 days window) obtained by subtracting the running mean of predetermined width, the deseasonalizing bin, from the original time series. This approach was also used in Emmert et al. (2012) resulting in the larger trend than obtained by Qian et al. (2017) in ACE-FTS CO_x (CO_2+CO), albeit on a time series 1 year shorter than used in Qian et al. (2017). The differences were argued to be due to the procedure of removing outliers (Qian et al., 2017). On the other hand, the analysis of the SABER CO_2 VMR by Yue et al. (2015) relied on a different method, in which the multiple linear regression (MLR) is applied directly to the CO_2 time series, although the original paper mistakenly stated the bin size of 60 days; in fact, they used the 30-day window. Without further evidence, it is not obvious which method provides a more accurate trend determination, or whether and how the mean of the residuals approach removes seasonal, local time, and latitudinal variation in a better way, as summarized in Qian et al. (2017). The methods are compared and discussed in Appendix A. We found that, at least for the SABER data set, the use of one or another method is insignificant compared to the role played by the size of the time averaging bin.

A great deal of work has been devoted to develop methods for trends characterization, albeit there is not a general consensus on the best. For convenience of intercomparisons, the secular trend is usually defined as a *linear trend* over a given time period. The time series is assumed to be separable as

$$y(t) = L(t) + r(t), \tag{1}$$

where $r(t)$ is a residual, typically containing other physically determined periodic variations along with a random noise component, all oscillating around the straight line $L(t) = a_0 + a_1 t$. A MLR is usually carried out fitting the observations with a parameterized models of the deterministically defined components simultaneously. A typical procedure in time series analysis also involves a preprocessing step, usually called deseasonalizing. This may involve, among others, treatment of outliers, data transformation, filtering, smoothing, and time series characterization such as time lags or variance determination.

In this study, we aim at understanding the reported discrepancies for the CO_2 SABER Level2C trends estimated by Qian et al. (2017) versus Yue et al. (2015). In particular, we address whether the different methodologies of trend estimation can explain the differences in trends and assess the role played by the satellite sampling. We focus only on the SABER CO_2 data set available from daytime observations, labeled as Level2C and described in section 2. We use SD-WACCM output, essentially *forward modeling* the observation's sampling of latitude, local time, and seasons, including the effects of missing data. The synthetic time series is formed from the SD-WACCM calculations sampled at the exact coordinates of SABER observations. In this way, as we control the sampling in the model data, we can assess its impacts on the trend and hopefully to find a way to correct for them in the SABER data.

Section 2 provides details of SABER data Level2C, including spatial and temporal sampling, and describes how the SD-WACCM model can be used to validate our trend methodology. The analysis of the SABER and SD-WACCM data is described in section 3. Section 4 summarizes our findings and provides discussion of the implications.

2. SD-WACCM and SABER Data

2.1. SD-WACCM Data set

The SD-WACCM developed at NCAR (Marsh et al., 2013) is an essential tool in this study. First, a linear trend estimated from the model output sampled at uniform time intervals, with no missing data, serves as a *truth* for subsequent comparisons. Second, the model is used to generate a synthetic time series of data sampled at the SABER Level2C CO_2 geolocations, on which we can test the performance of the different procedures of trend estimation and deseasonalizing window widths. As already shown in Rezac, Jian, et al. (2015), the global mean vertical profile of CO_2 VMR derived from SABER compares significantly better with SD-WACCM data when sampled at SABER geolocations than when using all SD-WACCM data. This is understood to be due to the nonuniform sampling of latitudes and local times by the SABER daytime data set.

An updated review of the processes that influence the vertical profile of CO_2 distribution is provided in Garcia et al. (2014), who also discussed the roles of turbulent versus diffusive mixing and pointed out the limited importance of photolysis below 110 km. We use the same SD-WACCM output, obtained with a Prandtl number, Pr = 4, as in Yue et al. (2015). A smaller Pr implies a stronger eddy diffusion in the model, which actually leads to a better agreement with the ACE-FTS CO_2 vertical profile if set to Pr = 2 (Garcia et al., 2014). However, this is not a significant issue in our analysis. Additional details and validation of the SD-WACCM physics related to the MLT, especially in connection to CO_2, can be found in Smith et al. (2011), later validated with SABER data in Rezac, Jian, et al. (2015) and recently by MIPAS data in López-Puertas et al. (2017).

2.2. SABER Data set Level2C

The SABER instrument description and characteristics, including satellite orbit and the viewing direction, are originally described in Russell et al. (1999), and also briefly reviewed in works already mentioned (e.g., Qian et al., 2017; Rezac, Kutepov, et al., 2015; Yue et al., 2015). The SABER CO_2 VMR data are post-operationally retrieved simultaneously with temperatures, however, only for daytime conditions (solar zenith angle <80°) as detailed in Rezac, Kutepov, et al. (2015). This data set is labeled Level2C and it is being continuously processed with new observations as they become available (The data can be found at http://saber.gats-inc.com/). In this analysis we use a 15-year long data set (2002–2016). Despite the daytime only retrieval restriction, the data provide a good latitude coverage (−54 to 54°), albeit nonuniformly sampled and affected by data gaps that drift from year to year (due to drift in the satellite's orbit; Sam Yee 2017, private communication). Moreover, the local time sampling for the same day-of-the-year is slightly drifting. These two effects are illustrated in Figure 1. Each data point represents a zonal mean of daytime CO_2 VMR at 100-km altitude and latitude (5.0±5°S). The data are plotted as a function of day-of-the-year, with different panels corresponding to different years, as labeled.

From Figure 2 we also see that the SABER daytime data are taken either in the morning or evening. Local noon is never sampled because SABER cannot look directly to the Sun (Figure 2). In addition, there are data gaps that occur during periods when daytime sampling is not possible (solar zenith angle is >80° for both ascending and descending orbits). The gaps shift about 1 day per year, hence about 15 days over the entire time series analyzed here (2002–2016). Furthermore, the local time sampling is a function of latitude as shown in Figure 2. Essentially, the Level2C data at a given latitude are sampled as a bimodal distribution of local times, the morning (6–11 hr) and the evening (13–18 hr). This bimodal distribution, however, remains nearly unchanged from year to year.

The shifting data gaps and local time sampling typical for a long satellite mission duration are expected to make an accurate determination of decadal trends more challenging. Other factors that impact the trend determination are the variations in the calibration of the instrument over time and the proper characterization of the natural variability of the atmosphere. Because trend detection is fundamentally determining a signal (the trend) above the *noise* (the natural variability), the length of the data set also plays a role in the ability to determine a trend to a specific uncertainty (Leroy et al., 2008). In addition, errors associated with incomplete spatial or temporal sampling due to the specific satellite orbit are also extremely important to consider. Furthermore, changes in the satellite orbit over time may result in the derivation of false trends. Wielicki

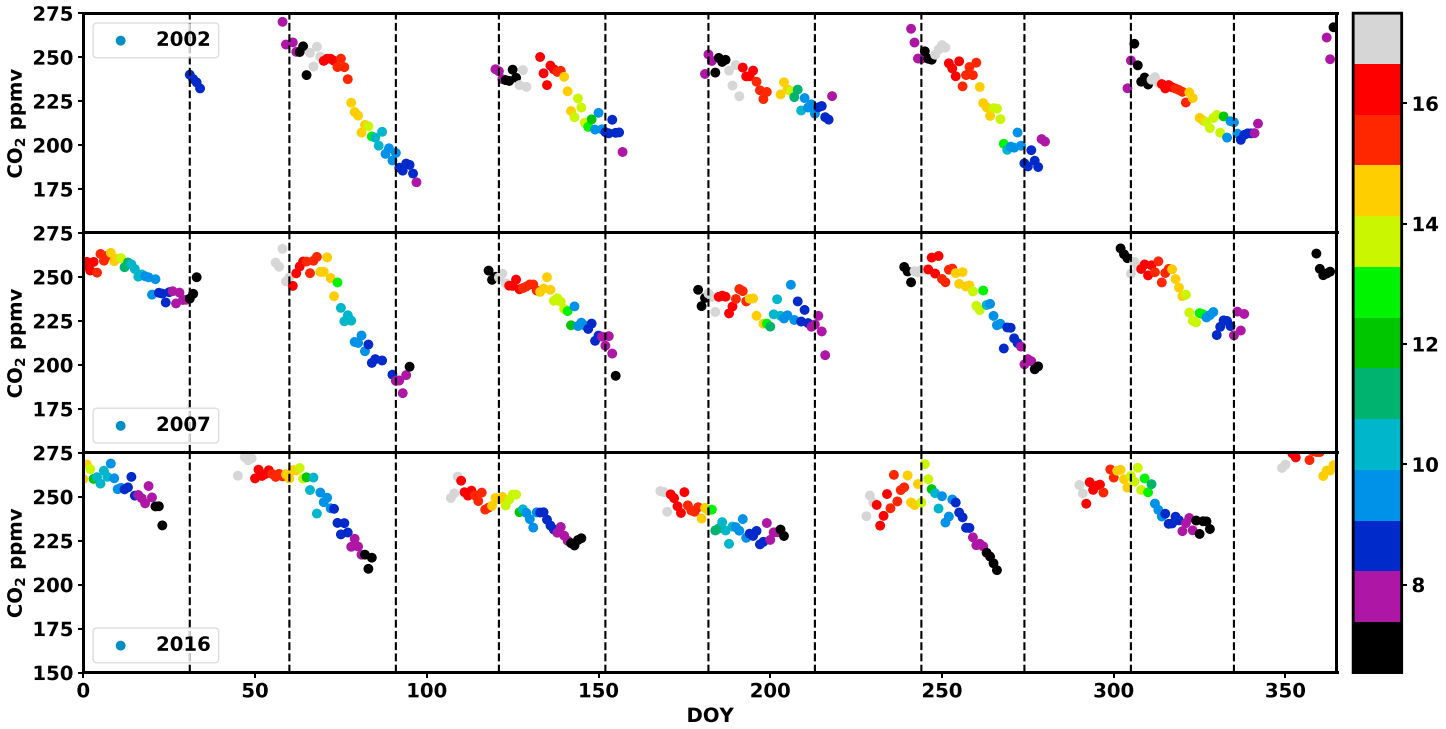

Figure 1. An example of Sounding of the Atmosphere using Broadband Emission Radiometry Level2C data coverage of the daytime zonal mean CO_2 volume mixing ratio at 100 km, averaged over latitude 0–10°S, demonstrating the data gaps shifting for a particular DOY over the years; sampled as 2002, 2007, and 2016 from top to bottom panels. The color bar indicates the average local time in hours for the data point (see text for a discussion). DOY = day-of-year.

et al. (2013) provide a comprehensive review of these error mechanisms related to the detection of trends in tropospheric climate parameters, although the discussion is generic to all trend detection from satellite observations.

3. Trends Analysis Results

In this work we follow the same methodology of time series analysis applied in Yue et al. (2015). We apply the MLR directly to the time series, where each data point represents a zonal mean within a specific time, latitude, and altitude (pressure) bin. The width of the temporal bin is the key subject of this analysis. In particular, we try to find out if the different (30- versus 60-day) bins yield a different trend. We also want to test whether the deseasonalizing method plays a significant role in the trend estimation as Qian et al. (2017) suggest (see Appendix A). The linear trends in this study are converted to *relative* trends (percent per decade) by dividing the estimated trend value by the constant term in $L(t)$. The deterministic part of the residual, $r(t)$ in equation (1), can be written as

$$r(t) = a_2 \cdot SAO + a_3 \cdot AO + a_4 \cdot QBO + a_5 \cdot Sol. \quad (2)$$

The semiannual (SAO) and annual oscillation (AO) are modeled as purely periodic waves. The QBO represents the quasi-biennial-oscillation approximated by the 30 mb zonal mean winds at the equator (http://www.geo.fu-berlin.de/met/ag/strat/produkte/qbo/qbo.dat), while the solar cycle term, *Sol*, is parametrized with the 10.7 cm radio flux (http://lasp.colorado.edu/lisird/tss/). As has been reported in Yue et al. (2015), the QBO terms can be neglected without significantly impacting the CO_2 linear trend determination.

We perform the analysis in isobaric coordinates for both WACCM and SABER to simplify interpretation of the results. Trends in CO_2 (which has a strong vertical gradient above 80 km) will be affected by contraction or expansion of the atmosphere when computed in geometric coordinates. Figure 3 shows the estimated trend using the global SD-WACCM output, WACCM-full, using 30- and 60-day bins. The latitude binning is 10°. There are no inherent data gaps in the SD-WACCM data set, and the sampling is uniform in latitude (same number of data points per bin).

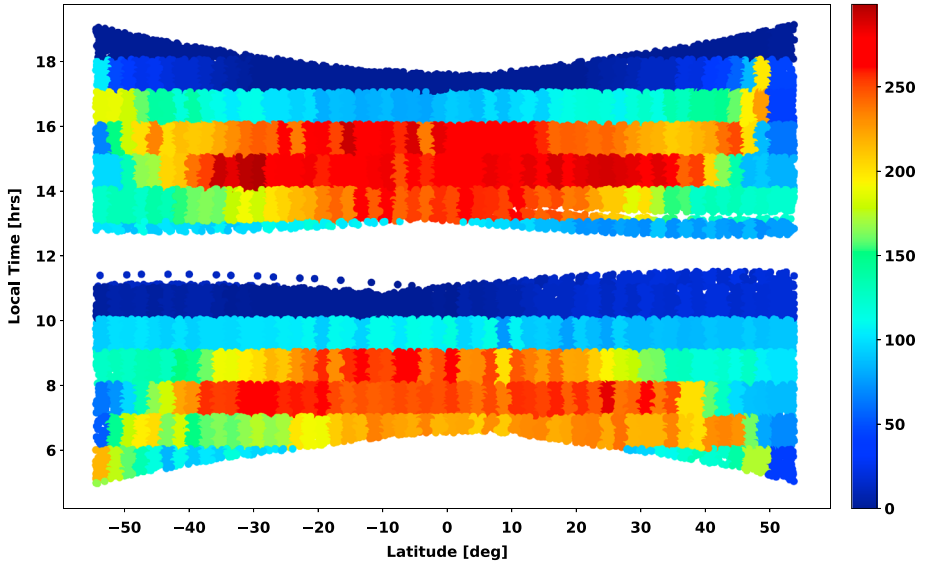

Figure 2. Latitude versus local time frequency distribution for the Sounding of the Atmosphere using Broadband Emission Radiometry Level2C data for 2003. The colorbar indicates the data count within the latitude, local time bin. In general, this distribution does not change significantly from year to year, neither overall over the 15-year period. This figure demonstrates that there are only two local time bands of sampling.

The WACCM-full results for the 30- and 60-day time bins indicate a very similar trend distribution in the latitude/pressure space. The minima and maxima locations are consistent in magnitude and location, with differences smaller than 1% per decade. Qualitatively we conclude that in general the 60-day binned data yield a slightly smaller trend above 90 km, although this difference is not statistically significant.

In the next step, we resample the original WACCM-full output to the SABER CO_2 geolocations. The results are shown in Figure 4. In this case the two time averaging windows yield somewhat different trend distributions with latitude. As already noted, the bimonthly averaging provides slightly smaller trends above about 90 km altitude, although there are additional localized differences below this height, typically smaller than the uncer-

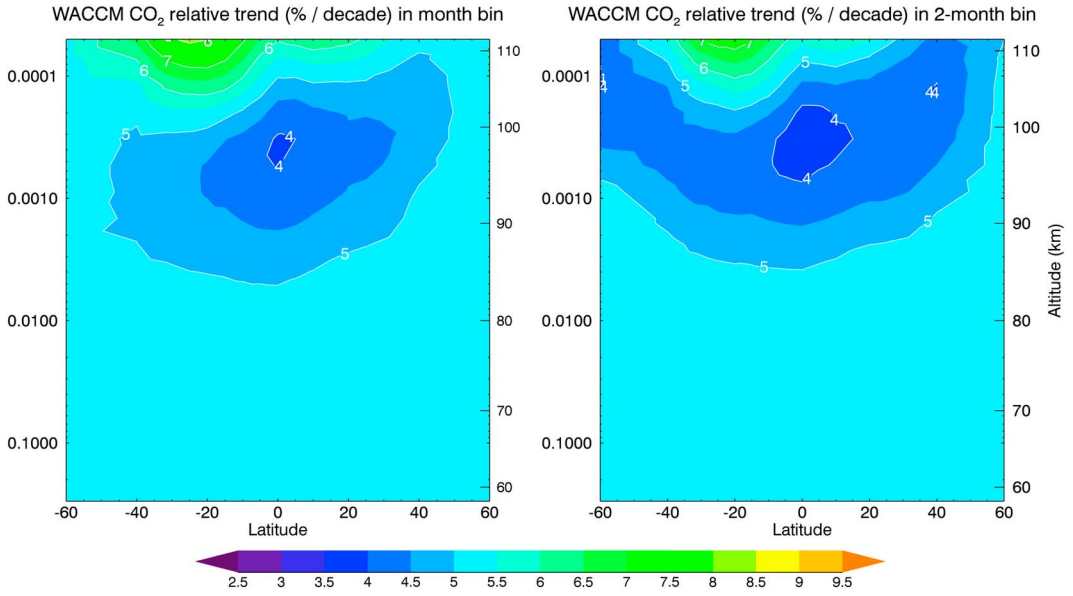

Figure 3. Secular (relative) trend in percent per decade of the WACCM-full output. (Left) estimated using the monthly (30-day) bin and (Right) using the bimonthly (60-day) bin. The pressure coordinates [mb] are used for the analysis and for plotting, with an approximate altitude scale shown on the right hand *y* axis. WACCM = Whole Atmosphere Community Climate Model.

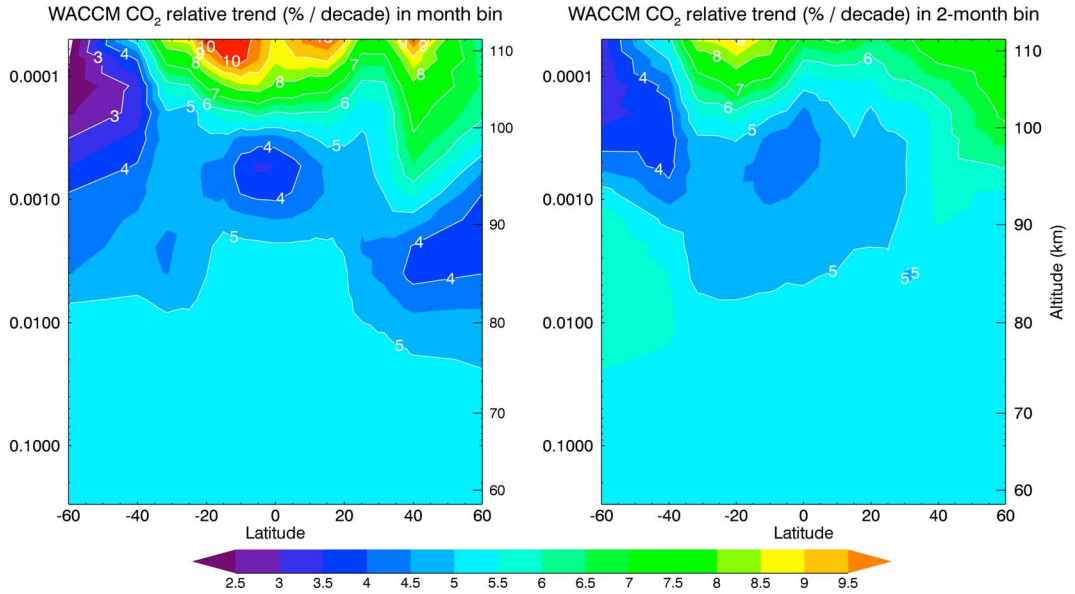

Figure 4. Secular trend in percent per decade of the SD-WACCM output sampled only at Sounding of the Atmosphere using Broadband Emission Radiometry geolocations (WACCM-sub). (Left) secular trend for the monthly (30-day) bin, and (right) using the bimonthly (60-day) bin. WACCM = Whole Atmosphere Community Climate Model.

tainty in the trend estimate (e.g., see Figure 5). Overall, the 30-day bin data results in noticeably larger trends above approximately 100 km for most of latitudes north of −40° (little less than 2% per decade).

Comparing the 30-day averages, left panels in Figure 3 (WACCM-full) with those in Figure 4 (WACCM-sub), we observe markedly different latitude-altitude distributions of the trends as well as absolute values. The WACCM-sub, especially above 100 km, yields trends that are nearly 2% per decade larger than those from WACCM-full. In addition, there is an overestimation of trends in the northern hemisphere (30–60°) above 95 km, and underestimation between 85 and 95 km relative to the WACCM-full results. Locally, the differences can be as large as 4% per decade, such as the peak near 40° at 105–110 km range (to some extend also at latitudes 10S, 10N). Similarly, comparing the 60-day averages (right panels of the respective figures), the WACCM-sub shows larger trends above 100 km, with differences growing with increasing latitudes from 30 to 60°. Overall, there is a better agreement in the distributions in the southern hemisphere, although there is a region of latitudes (~45–60S) above 100 km altitude where the WACCM-sub yields up to 1% per decade smaller values than the trends from WACCM-full. The 60-day time window averages produce maximum differences just over 3% per decade in the higher northern latitudes.

The trend analysis of the SABER data is shown in Figure 6. Averaging data in 30-day bins (left panel) leads to a linear trend which is substantially larger than that derived from the 60-day averaging (right panel) at high altitudes (>95 km). The trend enhancement for the 30-day binning is largest in the latitude range 40S–40N, with maximum values as large as 10–14% per decade. The differences between the trends computed using 30- versus 60-day binned data in both WACCM and SABER suggest that tidal aliasing due to incomplete and shifting local time sampling might account for the large trends obtained using 30-day binned data. The differences between 30- and 60-day bin cross-sections are generally consistent with the results obtained using the SD-WACCM data, although the maximum values of the differences are noticeably larger for SABER. This might be due to inaccuracies in the simulation of the tides in WACCM, if indeed tidal aliasing is involved. We are performing another designated study on the tides in CO_2 VMR, which is outside of the scope of this work. It seems that in the 60-day binning local time samples occur more uniformly compared to the 30-day binning, thus tending to remove the tidal effects, at least for the period under investigation (2002–2016). In addition, it is clear that the hemispheric asymmetry found by Yue et al. (2015) is caused by the use of the 30-day averaging window. This asymmetry is largely removed with the 60-day bins.

Some of these key points can be better illustrated with the global mean vertical profiles of the relative trends (in pressure coordinates) shown in Figure 5.

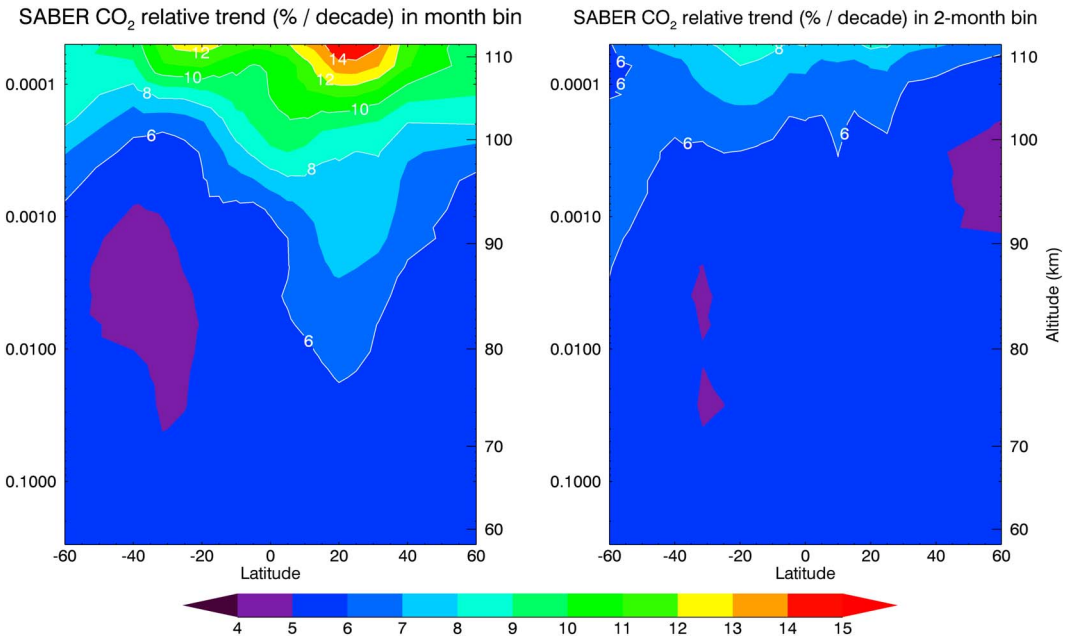

Figure 5. Global mean vertical profile of relative trends derived from WACCM-full (black curve), WACCM-sub (red), and SABER data (blue). (left) 30- and (middle) 60-day averaging bins and their differences [percent/decade] (right panel) are shown. The 1σ uncertainties are shown as shaded areas. WACCM = Whole Atmosphere Community Climate Model; SABER = Sounding of the Atmosphere using Broadband Emission Radiometry.

From the global mean differences (30- versus 60-day window) in the right panel in Figure 5, the SABER monthly mean data lead to overestimation of the relative trend by about 3% per decade above 1×10^{-3} mb (~93 km). For the WACCM-full the differences are less than 1% per decade, while for WACCM-sub they are above 1% per decade around 1×104^{-4} mb. Importantly, the bimonthly time averages always produce smaller trends (above 1×10^{-3} mb), possibly due to different local time variations in SD-WACCM CO_2 VMR. Detailed diagnosis of this factor will be performed in a future study. In addition, there is close resemblance between the vertical profile shapes for the SABER data (blue) and WACCM-sub (red), although the absolute values do not agree. Furthermore, we should also note that the SABER bimonthly relative trend is about 8% per decade around

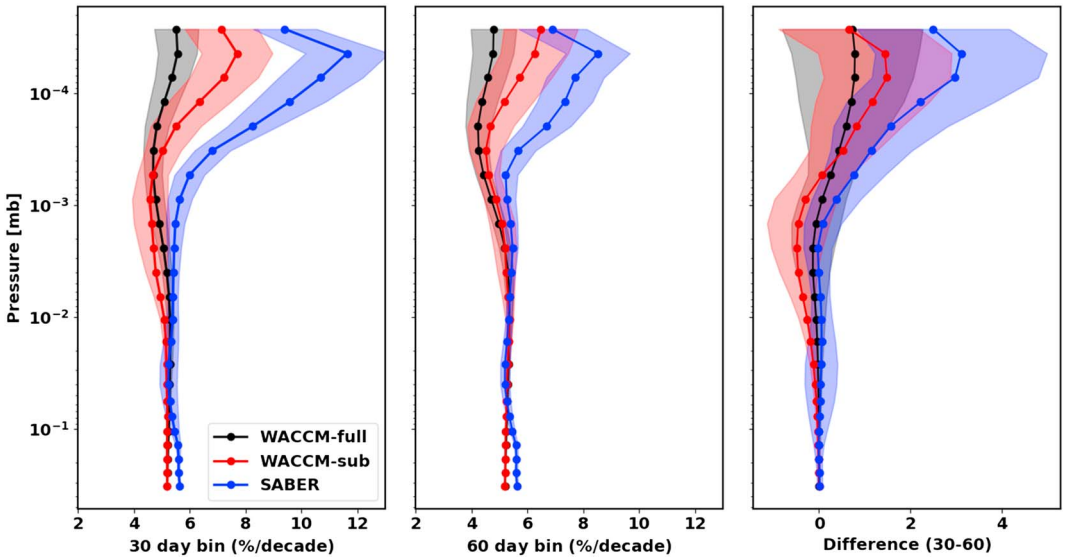

Figure 6. Secular trend in percent per decade derived for the SABER CO_2 volume mixing ratio data set Level2C. (Left) secular trend for the monthly bin and (Right) using the bimonthly bin. SABER = Sounding of the Atmosphere using Broadband Emission Radiometry.

1×10^{-4} mb (110 km), which is consistent (within error bars) with those derived in Qian et al. (2017). We would like to stress that the agreement between our results, when accounting for the 60-day binning, and Qian et al. (2017) rules out the *deseasonalization* as the cause of the discrepancy, as claimed in Qian et al. (2017; (see Appendix A). The 60-day binning removes spurious trends caused by the nonuniformity of the local time sampling over the considered period, as can be appreciated from Figure 1. These data can still be aliased by the tide (nighttime data are not yet available), but the aliasing is uniform and should not affect the trends, unless there is a trend in the tides themselves.

4. Discussion and Conclusions

We have investigated the effects of sampling in the SABER daytime CO_2 data (Level2C) in regards to the linear trend estimation, along with the effects of time window averaging, for example, monthly and bimonthly averages. A standard MLR analysis was applied to extract the linear trend from the time series data. In order to better isolate the effects of temporal sampling due to data gaps shift with years, and changing time window averaging, we use the SD-WACCM data in a *forward model* analogy. Specifically, the SD-WACCM output was used to generate two time series: the WACCM-full, where the latitude and time sampling are uniform, with no data gaps and has same number of data points in each latitude/time bin; WACCM-sub where SD-WACCM was sampled at SABER geolocations. In addition, we also reanalyze the SABER data by using monthly and bimonthly time averages.

The daytime CO_2 SABER data suffer from several sampling biases. First, the data is typically sampled at local times either in the morning (6–11 hr) or evening (13–18 hr). In addition, because the yaw cycle imposed on the spacecraft undergoes a progressive drift, there is also a shift in local time for a given location over the 16-year period of SABER observations. Similarly, the data gaps shift on the order of 1 day per year at low latitudes, but can be 2–3 for certain years in latitudes near 50S or 50N.

We have demonstrated that the monthly (30-day) bin time series leads to linear trends significantly larger than for the bimonthly (60-day) averaged data. The bimonthly bin averages are better suited for the linear trend analysis of CO_2 SABER since each yaw-cycle phase of the spacecraft takes about 60 days. In this way the number of data points and the local time sampling remains approximately the same within the bin from year to year. The 60-day window also helps to mitigate the data gap problem (and their shift). Each period of missing SABER data is around 20 days and it can pose more severe problems when using a monthly bin. This interpretation is also consistent from the analysis of the synthetic time series generated from the SD-WACCM data. In general we can summarize the main points as follows:

- The SABER Level2C data set has a specific pattern of missing data and local time sampling both of which drift with time due to the orbit precession and yaw-cycle drift. We have shown by analysis of this data set that using monthly averages lead to an overestimated linear trend, while the 60-day means can alleviate this problem by providing a more uniform local time distribution. If the SABER CO_2 VMR did not depend on local time, the fact that alternate yaw cycles are poorly sampled, or that local time drifts, would not affect the calculation of the linear trend. The most likely source of the CO_2 VMR variability with local time is the tides above 80 km (where the tides have large amplitudes and the CO_2 has a strong vertical gradient).
- The SABER global mean relative trend from the bimonthly time averaged data is consistent with the tropospheric value up to ≈90 km altitude. The vertical profile shows a small increase with altitude above 1×10^{-3} mb (~90 km), reaching a maximum value of 8.2% per decade at approx. 1×10^{-4} mb (105–110 km). This deviation from the tropospheric value is only marginally statistically significant. These results are also consistent with the ACE-FTS data presented in Emmert et al. (2012).
- The SD-WACCM data do not show any tendency of increasing trend with altitude (above the 1×10^{-3} mb [~90 km]) whereas SABER data do show such behavior. So far no explanation has been proposed for this except the change of the transport or eddy diffusion (Emmert et al., 2012; Garcia et al., 2016).
- This reanalysis of the SABER CO_2 data agrees well with the results of Qian et al. (2017) in the absolute value of the global mean relative trend (both performed on the pressure coordinate). However, we conclude that the key choice is only the proper time averaging bin, rather than selecting a specific deseasonalizing method as hypothesized in Qian et al. (2017). This conclusion was already hinted on in the Figure 3 of their work (where residuals are always averaged in 60-day bins, and only deseasonalizing window is varied; except their blue curve). We have demonstrated that the deseasonalization has a negligible effect (see Appendix A).

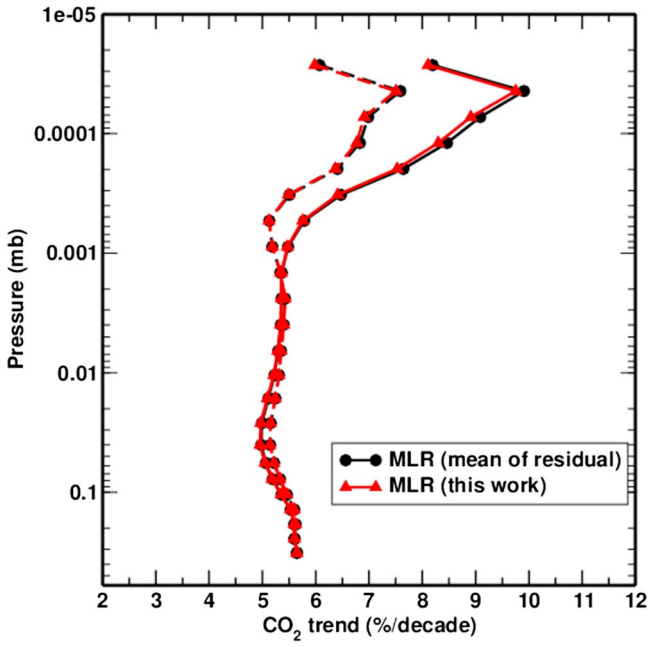

Figure A1. Relative trends derived with the *mean of residual* (lines with circle symbols) method as described in Qian et al. (2017) and a MLR applied directly to the SABER CO_2 data as used in this work (lines with triangle symbols). The dashed lines represent the bimonthly time bins, while the full lines show the trend derived using monthly binning. Both methodologies provide very consistent estimates for a given time averaging bin. MLR = multiple linear regression.

- We recommend that the daytime SABER CO_2 Level2C data set should be averaged in time windows of 60 days for long-term trend studies. Such averaging provides more uniform local time sampling, and reduces aliasing of diurnal tides into the long-term trend.

If the SABER instrument remains healthy, it would be a significant benefit if operations could be extended to cover possibly two solar cycles to obtain more robust trend estimates. Another significant improvement, not only for trend studies, would be to retrieve CO_2 VMR from the SABER nighttime measurements to fill in most of the spatial gaps and local time sampling. While this is desirable, it is also a difficult task because of the combined effects of low signal-to-noise at night and non-LTE that must be included in the radiative transfer (Panka et al., 2017). The implications from studying the links and effects of increasing CO_2 in the troposphere on the upper atmosphere holds clues not only for the immediate present or future, but also understanding of its past.

Appendix A: MLR methodology trend comparisons

Estimating the secular changes from the satellite data set, such as SABER, requires removal of all other *natural variations*, both spatial and temporal ones. This, in principle, also encompasses possible errors due to orbit and sampling changes. In Qian et al. (2017) two methodologies are described to accomplish this, referred to as the *deseasonalizing* step. Their conclusion is that the method of *mean of residual* is somewhat superior as it shows a weaker dependence on the deseasonalizing bin size, although the evidence provided (their Figure 3) does not show the effects of varying the averaging window (for all the cases). Here we compare SABER CO_2 trends estimated using the mean of residual method and the approach of applying the MLR directly to the data using explicit terms, including the annual and semiannual variations in the fit model (as in equation (2)). The results are presented in Figure A1. The dashed set of curves are for the bimonthly averaged time series, while the full curves correspond to the monthly binning. First, we note that for a given time averaging window both methods provide a consistent trend estimation. Second, the differences in trends are significantly smaller between the two methodologies, than the differences due to varying the time averaging window (30- versus 60-day). Therefore, for the reasons discussed in the main text, a proper time averaging of the SABER data is much more important than a particular methodology applied.

Acknowledgments
The authors would like to acknowledge the support of the SABER retrieval team providing the version 2.0 data, including scientists from GATS, Inc., NASA Langley Research Center, NASA Goddard Space Flight Center, Spain (IAA), and Arcon, Inc. The SABER CO_2 data used in this study are available to the public in the form of NetCDF files at ftp://saber.gats-inc.com/Version2_0/Level2C/. More information about WACCM can be found at https://www2.acom.ucar.edu/gcm/waccm. The SD-WACCM outputs are available on NCAR's High Performance Storage System. The MERRA data can be obtained from http://disc.sci.gsfc.nasa.gov/daac-bin/DataHoldings.pl. This research was carried out at the Hampton University and Max-Planck-Institut für Sonnensystemforschung. L. R. acknowledges financial support from the Priority Program 1788 *Dynamic Earth* of the German Research Foundation (DFG). J. Y. is supported by NASA grants NNX14AF20G, NNH13ZDA001N-HGI, and NNH15ZDA001N-HSR. J. Y. is grateful to Liying Qian, Quan Gan, and Anne Smith for valuable discussions. J. Y. is also supported by the ISSI team "Climate Change in the Upper Atmosphere". M. L.-P. has been supported by the Spanish MICINN under projects ESP2014-54362-P and ESP2017-87143-R and EC FEDER funds.

References

Beagley, S. R., Boone, C. D., Fomichev, V. I., Jin, J. J., Semeniuk, K., McConnell, J. C., & Bernath, P. F. (2010). First multi-year occultation observations of CO_2 in the MLT by ACE satellite: Observations and analysis using the extended CMAM. *Atmospheric Chemistry and Physics*, *10*(3), 1133–1153. https://doi.org/www.atmos-chem-phys.net/10/1133/2010/

Bernath, P. F., McElroy, C. T., Abrams, M. C., Boone, C. D., Butler, M., Camy-Peyret, C., et al. (2005). Atmospheric chemistry experiment (ACE): Mission overview. *Geophysical Research Letters*, *32*, L15S01. https://doi.org/10.1029/2005GL022386

Emmert, J. T. (2009). A long-term data set of globally averaged thermospheric total mass density. *Journal of Geophysical Research*, *114*, A06315. https://doi.org/10.1029/2009JA014102

Emmert, J. T. (2015). Altitude and solar activity dependence of 1967–2005 thermospheric density trends derived from orbital drag. *Journal of Geophysical Research: Space Physics*, *120*, 2940–2950. https://doi.org/10.1002/2015JA021047

Emmert, J. T., Stevens, M. H., Bernath, P. F, Drob, D. , & Boone, C. D. (2012). Observations of increasing carbon dioxide concentration in Earth's thermosphere. *Nature Geoscience*, *5*, 868–871. https://doi.org/10.1038/ngeo1626

Garcia, R. R., López-Puertas, M., Funke, B., Kinnison, D. E., Marsh, D. R., & Qian, L. (2016). On the secular trend of CO_x and CO_2 in the lower thermosphere. *Journal of Geophysical Research: Atmospheres*, *121*, 3634–3644. https://doi.org/10.1002/2015JD024553

Garcia, R. R., López-Puertas, M., Funke, B., Marsh, D. R., Kinnison, D. E., Smith, A. K., & González-Galindo, F. (2014). On the distribution of CO_2 and CO in the mesosphere and lower thermosphere. *Journal of Geophysical Research: Atmospheres*, *119*, 5700–5718. https://doi.org/10.1002/2013JD021208

Laštovička, J., Beig, G., & Marsh, D. R. (2014). Response of the mesosphere-thermosphere-ionosphere system to global change - CAWSES-II contribution. *Progress in Earth and Planetary Science*, *1*(1), 21. https://doi.org/10.1186/s40645-014-0021-6

Leroy, S. S., Anderson, J. G., & Ohring, G. (2008). Climate signal detection times and constraints on climate benchmark accuracy requirements. *Journal of Climate*, *21*, 841–846. https://doi.org.doi:10.1175/2007JCLI1946.1

López-Puertas, M., Funke, B., Jurado-Navarro, Á. A., García-Comas, M., Gardini, A., Boone, C. D., et al. (2017). Validation of the MIPAS CO_2 volume mixing ratio in the mesosphere and lower thermosphere and comparison with WACCM simulations. *Journal of Geophysical Research: Atmospheres*, *122*, 8345–8366. https://doi.org/10.1002/2017JD026805

Marsh, D. R., Mills, M. J., Kinnison, D. E., Lamarque, J. F., Calvo, N., & Polvani, L. M. (2013). Climate change from 1850 to 2005 simulated in ESM1(WACCM). *Journal of Climate*, *26*(19), 7372–7391. https:doi.org/10.1175/JCLI-D-12-00558.1

Panka, P. A., Kutepov, A. A., Kalogerakis, K. S., Janches, D., Russell, J. M., Rezac, L., et al. (2017). Resolving the mesospheric nighttime 4.3μm emission puzzle: Comparison of the $CO_2(\nu_3)$ and OH(ν) emission models. *Atmospheric Chemistry & Physics*, *17*, 9751–9760. https:doi.org/10.5194/acp-17-9751-2017

Qian, L., Burns, A. G., Solomon, S. C., & Wang, W. (2017). Carbon dioxide trends in the mesosphere and lower thermosphere. *Journal of Geophysical Research: Space Physics*, *122*, 4474–4488. https://doi.org/10.1002/2016JA023825

Qian, L., Marsh, D., Merkel, A., Solomon, S. C., & Roble, R. G. (2013). Effect of trends of middle atmosphere gases on the mesosphere and thermosphere. *Journal of Geophysical Research: Space Physics*, *118*, 3846–3855. https://doi.org/10.1002/jgra.50354

Rezac, L., Jian, Y., Yue, J., Russell, J. M., Kutepov, A., Garcia, R., et al. (2015). Validation of the global distribution of CO_2 volume mixing ratio in the mesosphere and lower thermosphere from SABER. *Journal of Geophysical Research: Atmospheres*, *120*, 12,067–12,081. https://doi.org/10.1002/2015JD023955

Rezac, L., Kutepov, A., Russell, J. M., Feofilov, A. G., Yue, J., & Goldberg, R. A. (2015). Simultaneous retrieval of T(p) and CO_2 VMR from two-channel non-LTE limb radiances and application to daytime SABER/TIMED measurements. *Journal of Atmospheric and Solar-Terrestrial Physics*, *130*, 23–42. https://doi.org/10.1016/j.jastp.2015.05.004

Roble, R. G., & Dickinson, R. E. (1989). How will changes in carbon dioxide and methane modify the mean structure of the mesosphere and thermosphere? *Geophysical Research Letters*, *16*(12), 1441–1444. https://doi.org/10.1029/GL016i012p01441

Russell, J. M., Mlynczak, M. G., Gordley, L. L., Tansock, J. J., & Esplin, R. W. (1999). Overview of the SABER experiment and preliminary calibration results. In A. M. Larar (Ed.), *Proc. SPIE 3756, Optical spectroscopic techniques and instrumentation for atmospheric and space research III*. Denver, CO: Society of Photo-Optical Instrumentation Engineers. https://doi.org/10.1117/12.366382

Smith, A. K., Garcia, R. R., Marsh, D. R., & Richter, J. H. (2011). WACCM simulations of the mean circulation and trace species transport in the winter mesosphere. *Journal of Geophysical Research*, *116*, D20115. https://doi.org/10.1029/2011JD016083

Wielicki, B. A., Young, D. F., Mlynczak, M. G., Thome, K. J., Leroy, S., Corliss, J., et al. (2013). Achieving climate change absolute accuracy in orbit. *Bulletin of the American Meteorological Society*, *94*(10), 1519–1539. https://doi.org/10.1175/BAMS-D-12-00149.1

Yue, J., Russell, J., Jian, Y., Rezac, L., Garcia, R., López-Puertas, M., & Mlynczak, M. G. (2015). Increasing carbon dioxide concentration in the upper atmosphere observed by SABER. *Geophysical Research Letters*, *42*, 7194–7199. https://doi.org/10.1002/2015GL064696

Zhang, S. R., & Holt, J. M. (2013). Long-term ionospheric cooling: Dependency on local time, season, solar activity, and geomagnetic activity. *Journal of Geophysical Research: Space Physics*, *118*, 3719–3730. https://doi.org/10.1002/jgra.50306